FDT3303 Critical Perspectives on Data Science and Machine Learning 7.5 credits

Kritiska perspektiv på datavetenskap och maskininlärning

This is a translation of the Swedish, legally binding, course syllabus.

Establishment

Course syllabus for FDT3303 valid from Autumn 2019

Grading scale

P, F

Education cycle

Third cycle

Specific prerequisites

The course can be taken by PhD students with sufficient experience in statistics, data science, and/or machine learning and artificial intelligence.

Language of instruction

The language of instruction is specified in the course offering information in the course catalogue.

Intended learning outcomes

Upon successful completion of this course, the student will be able to:

• describe and explain problems and pitfalls when interpreting standard experiments performed in these disciplines

• interpret existing work based on fundamental principles (e.g., no free lunch, bias-variance tradeoff, information theory, etc.)

• identify weaknesses and limitations of an existing work, and assess the claims made from the evidence presented

• analyse the reproducibility and replicability of an existing work

• assess the ethical implications of an existing work

• propose improvements to an existing work

Course contents

The main content of the course is through the presentation of a series of "classic" articles that critically reflect upon work in data science and machine learning, and related disciplines, e.g., applied statistics.

Disposition

The course consists of the following learning activities:

• Seminar sessions with group discussions

• Research paper presentations

• Homework as preparation for the seminars, which consists of reading assigned work and writing a critical annotation.

Course literature

Example literature for review include:

• W. Kimball, "Errors of the third kind in statistical consulting," J. American Statistical Assoc., vol. 52, pp. 133–142, June 1957.

• D. J. Hand, "Deconstructing statistical questions," J. Royal Statist. Soc. A (Statistics in Society), vol. 157, no. 3, pp. 317–356, 1994.

• D. J. Hand, "Classifier technology and the illusion of progress," Statistical Science, vol. 21, no. 1, pp. 1–15, 2006.

• K. L. Wagstaff, "Machine learning that matters," in Proc. Int. Conf. Machine Learning, pp. 529–536, 2012.

• C. Drummond and N. Japkowicz, "Warning: Statistical benchmarking is addictive. Kicking the habit in machine learning," J. Experimental Theoretical Artificial Intell., vol. 22, pp. 67–80, 2010.

• P. Langley, "Advice to authors of machine learning papers," Machine Learning, vol. 5, pp. 233–237, 1990.

• R. Holte, "Very simple classification rules perform well on most commonly used datasets," Machine Learning, vol. 11, pp. 63–91, 1993.

• E. Keogh and J. Lin, "Clustering of time series subsequences is meaningless: Implications for past and future research," in Knowledge and Information Systems, Springer-Verlag, 2004.

• E. R. Dougherty and L. A. Dalton, "Scientific knowledge is possible with small-sample classification," EURASIP J. Bioinformatics and Systems Biology, vol. 2013:10, 2013.

• J. Bryson and A. Winfield, "Standardizing ethical design for artificial intelligence and autonomous systems," Computer, vol. 50, pp. 116–119, May 2017.

• A.-L. Boulesteix, "Ten simple rules for reducing overoptimistic reporting in methodological computational research," PLoS Comput Biol, vol. 11, p. e1004191, 04 2015.

Equipment

None

Examination

 • EXA1 - Examination, 7.5 credits, grading scale: P, F

Based on recommendation from KTH's coordinator for disabilities, the examiner will decide how to adapt an examination for students with documented disability.

The examiner may apply another examination format when re-examining individual students.

If the course is discontinued, students may request to be examined during the following two academic years.

Examination includes a short research project that has to be documented in a written report and an oral presentation.

Other requirements for final grade

20 minute oral presentation during one seminar

80% of seminar preparations (homework)

Approved project report

Ethical approach

- All members of a group are responsible for the group's work.
- In any assessment, every student shall honestly disclose any help received and sources used.
- In an oral assessment, every student shall be able to present and answer questions about the entire assignment and solution.